

## **Supplementary materials**

### **Characterization of bile salt hydrolase from *Lactobacillus gasseri* FR4 and demonstration of its substrate specificity and inhibitory mechanism using molecular docking analysis**

Running Title: Bile salt hydrolase from *Lactobacillus gasseri* FR4

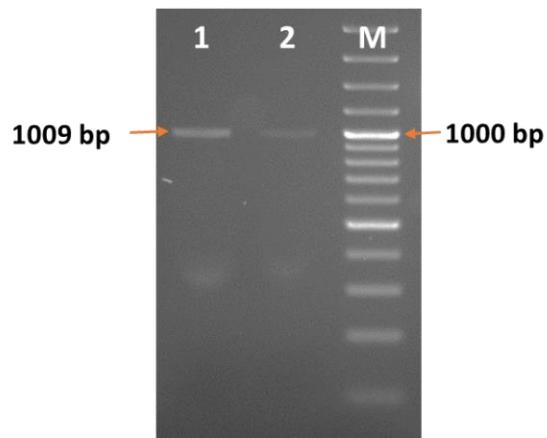
Rizwana Parveen Rani<sup>a</sup>, Marimuthu Anandharaj<sup>b</sup> and Abraham David Ravindran<sup>a\*</sup>

<sup>a</sup>Department of Biology, The Gandhigram Rural Institute-Deemed University, Gandhigram-624302, Tamilnadu, India.

<sup>b</sup>Biodiversity Research Center, Academia Sinica, Taipei-115, Taiwan.

E-Mail: rizm\_green@yahoo.com, anandharaj49@gmail.com and  
david\_gribiology@rediffmail.com

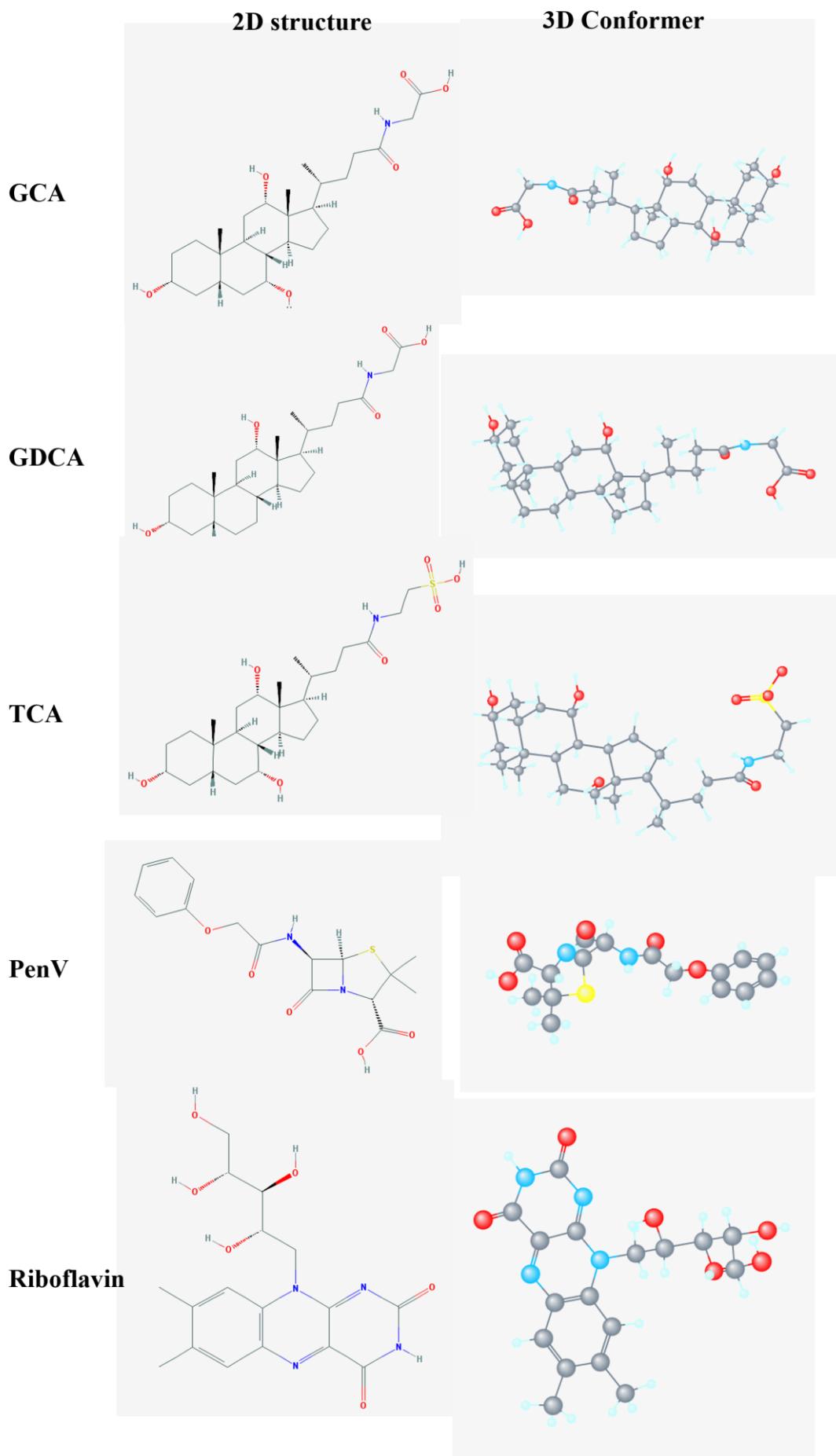
## Supplementary Figures



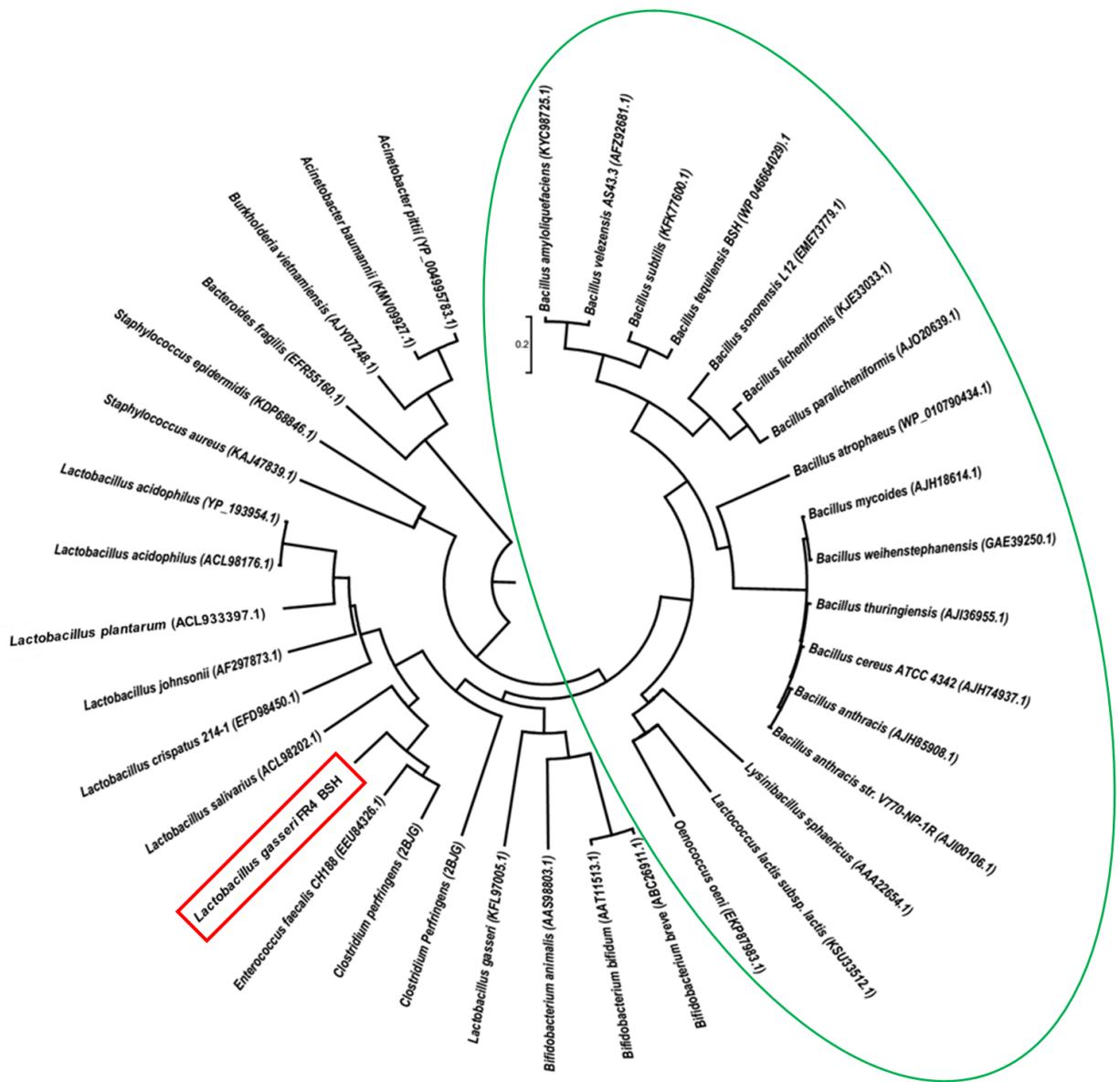
**Fig. S1:** Amplification of BSH gene from genomic DNA of *L. gasseri* FR4. Lane 1 and 2: BSH. M: Marker GeneRuler 100bp DNA ladder plus (Fermentas, USA).



**Fig. S2:** Sequencing results of *LgBSH* encoding gene using T7-R primer. The amplified *LgBSH* gene does not contain any mutations. The restriction enzymes *Xba*I and *Bam*HI were highlighted in green and red colour, respectively.



**Fig. S3:** 2D and 3D structures of ligands used in this study



**Fig. S4:** Phylogenetic analysis of *LgBSH* with BSH and PVA of various bacterial species was performed using UPGMA method. The PVA sequences were marked by green circle. The optimal tree with the sum of branch length = 9.23244762 is shown. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method and are in the units of the number of amino acid substitutions per site. The analysis involved 37 amino acid sequences. All positions containing gaps and missing data were eliminated. Evolutionary analyses were conducted in MEGA6 software.

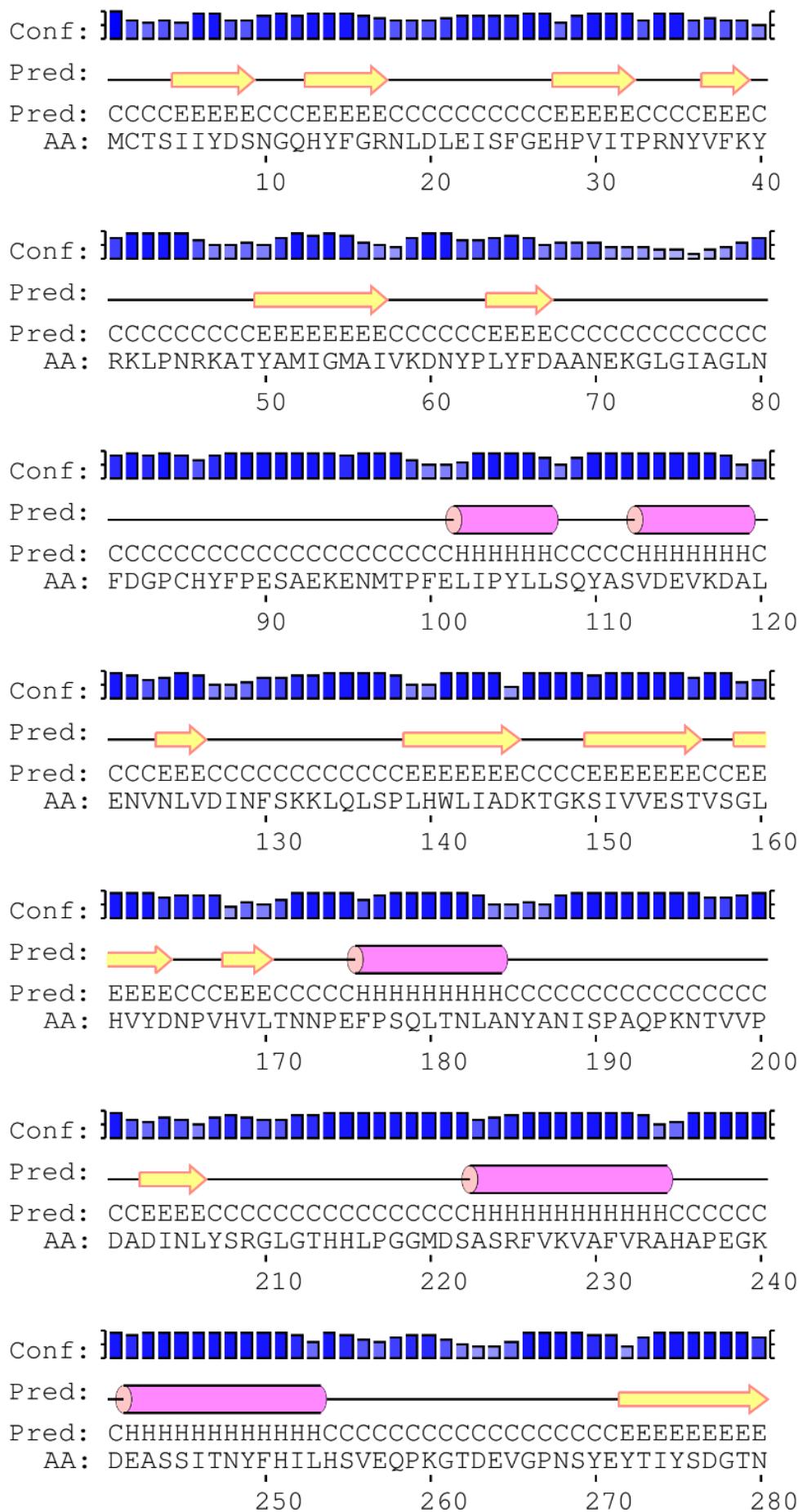
# Phyre<sup>2</sup>

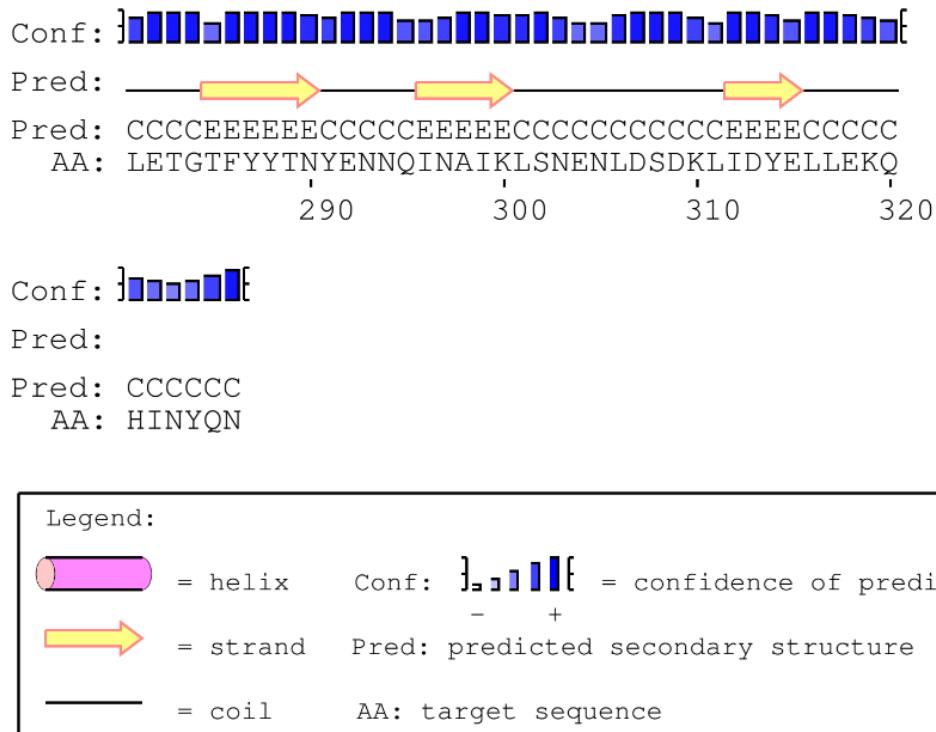
Email: anandharaj49@gate.sinica.edu.tw  
 Description: LgBSH  
 Date: Sun Sep 11 12:52:00 BST 2016  
 Unique Job ID: a73457768ec8d3f5

Detailed template information

#	Template	Alignment Coverage	3D Model	Confidence	% i.d.	Template Information
1	d2pvaa_	Alignment		100.0	30	<b>Fold:</b> Ntn hydrolase-like <b>Superfamily:</b> N-terminal nucleophile aminohydrolases (Ntn hydrolases) <b>Family:</b> Penicillin V acylase
2	c4wl3A_	Alignment		100.0	54	<b>PDB header:</b> hydrolase <b>Chain: A: PDB Molecule:</b> bile salt hydrolase; <b>PDBTitle:</b> crystal structure determination of bile salt hydrolase from2 enterococcus faecalis
3	c2bjgB_	Alignment		100.0	38	<b>PDB header:</b> hydrolase <b>Chain: B: PDB Molecule:</b> choloylglycine hydrolase; <b>PDBTitle:</b> crystal structure of conjugated bile acid hydrolase from2 clostridium perfringens in complex with reaction products3 taurine and deoxycholate
4	c5hkeB_	Alignment		100.0	47	<b>PDB header:</b> hydrolase <b>Chain: B: PDB Molecule:</b> bile salt hydrolase; <b>PDBTitle:</b> bile salt hydrolase from lactobacillus salivarius
5	c2hezB_	Alignment		100.0	37	<b>PDB header:</b> hydrolase <b>Chain: B: PDB Molecule:</b> bile salt hydrolase; <b>PDBTitle:</b> bifidobacterium longum bile salt hydrolase
6	c2oqcB_	Alignment		100.0	32	<b>PDB header:</b> hydrolase <b>Chain: B: PDB Molecule:</b> penicillin v acylase; <b>PDBTitle:</b> crystal structure of penicillin v acylase from bacillus subtilis
7	c4wl2F_	Alignment		100.0	22	<b>PDB header:</b> hydrolase <b>Chain: F: PDB Molecule:</b> putative exported choloylglycine hydrolase; <b>PDBTitle:</b> structure of penicillin v acylase from pectobacterium atrosepticum
8	c3hbcA_	Alignment		100.0	20	<b>PDB header:</b> hydrolase <b>Chain: A: PDB Molecule:</b> choloylglycine hydrolase; <b>PDBTitle:</b> crystal structure of choloylglycine hydrolase from bacteroides2 thetaiotaomicron vpi
9	c2x1cA_	Alignment		99.9	11	<b>PDB header:</b> transferase <b>Chain: A: PDB Molecule:</b> acyl-coenzyme <b>PDBTitle:</b> the crystal structure of precursor acyl coenzyme2 a:isopenicillin n acyltransferase from penicillium3 chrysogenum
10	c3gvzB_	Alignment		99.7	19	<b>PDB header:</b> structural genomics, unknown function <b>Chain: B: PDB Molecule:</b> uncharacterized protein cv2077; <b>PDBTitle:</b> crystal structure of the protein cv2077 from2 chromobacterium violaceum. northeast structural genomics3 consortium target cvr62
11	c1gk0D_	Alignment		83.5	21	<b>PDB header:</b> hydrolase <b>Chain: D: PDB Molecule:</b> cephalosporin acylase; <b>PDBTitle:</b> structure-based prediction of modifications in2 glutarylaminidase to allow single-step enzymatic production3 of 7-aminocephalosporanic acid from cephalosporin c

**Fig. S5:** Template selection using Phyre2 software. The BSH from *E. faecalis* (4WL3) was selected as a template for homology modeling.



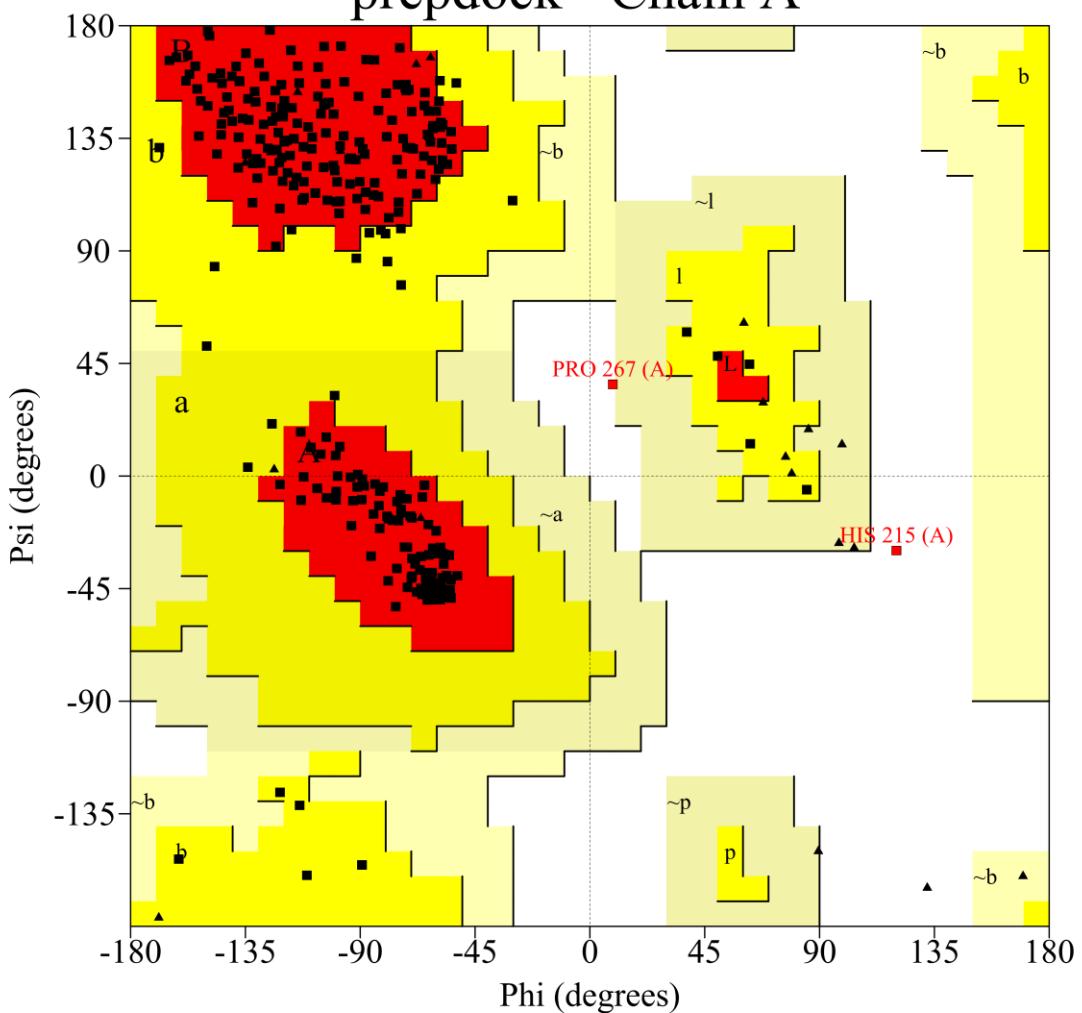


**Fig. S6.** Secondary structure prediction of *LgBSH* using PSIPRED V3.3.

PROCHECK

# Ramachandran Plot

prepdock - Chain A



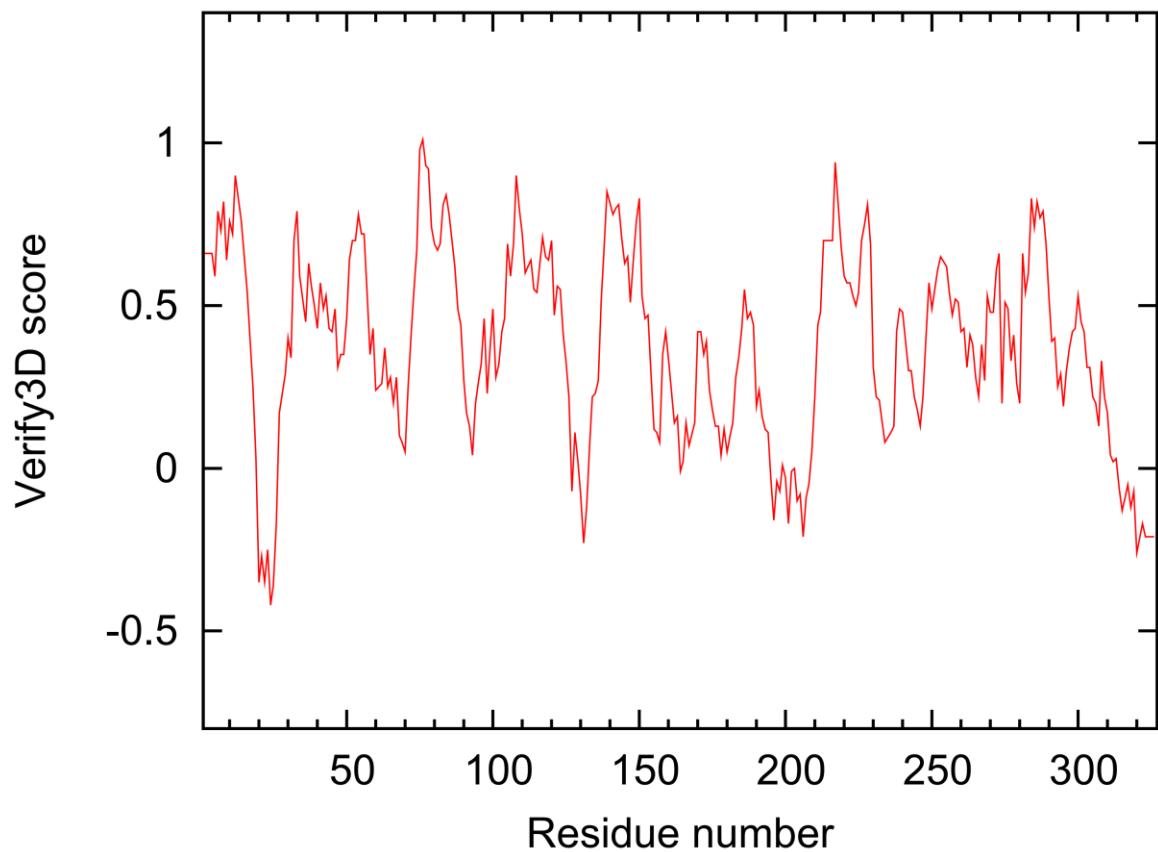
## Plot statistics

Residues in most favoured regions [A,B,L]	261	91.6%
Residues in additional allowed regions [a,b,l,p]	23	8.1%
Residues in generously allowed regions [~a,~b,~l,~p]	0	0.0%
Residues in disallowed regions	1	0.4%
Number of non-glycine and non-proline residues	285	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	19	
Number of proline residues	19	
Total number of residues	325	

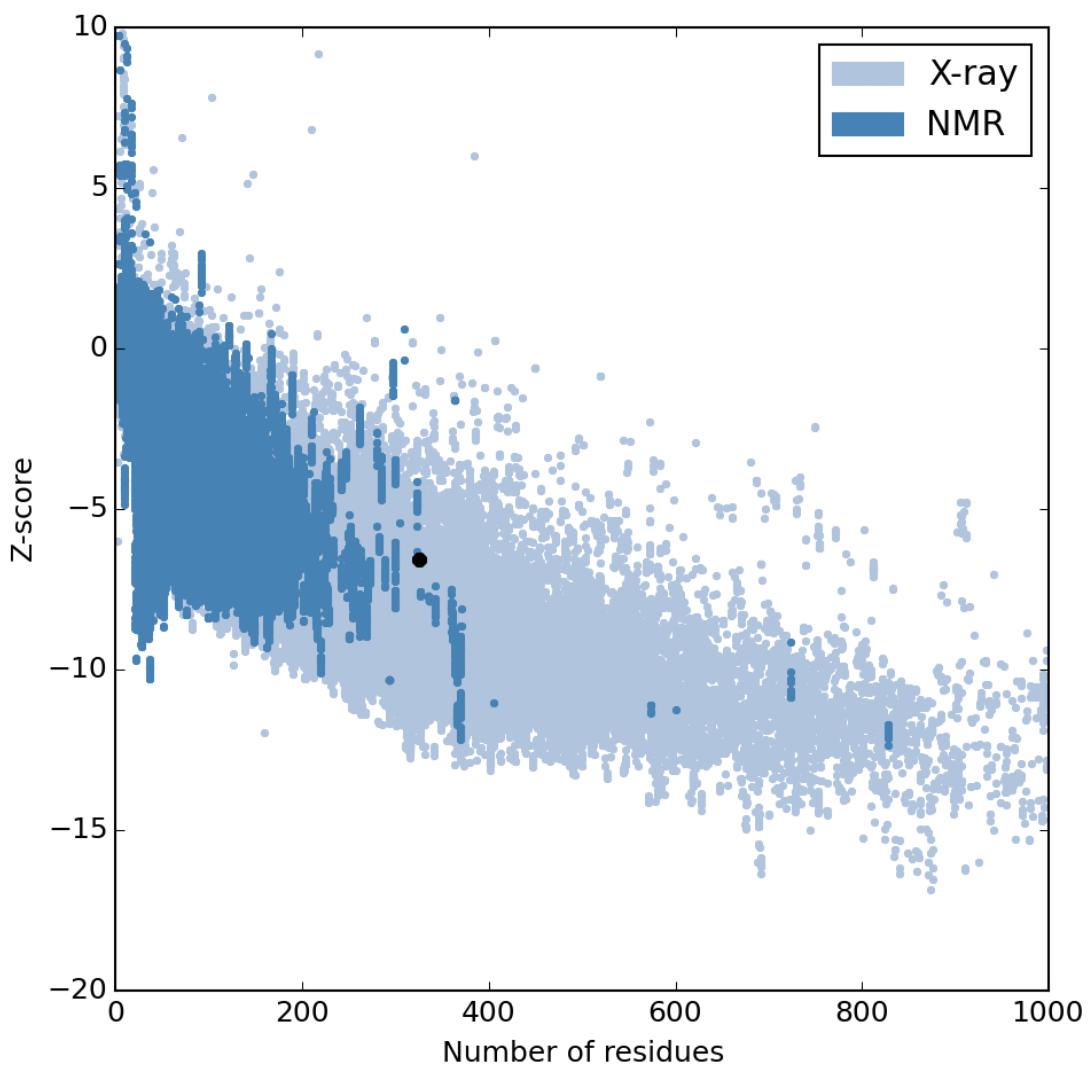
Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

**Fig. S7:** Ramachandran plot analysis of *LgBSH* using PSVS online software.

### Verify3D score over window of 7 residues

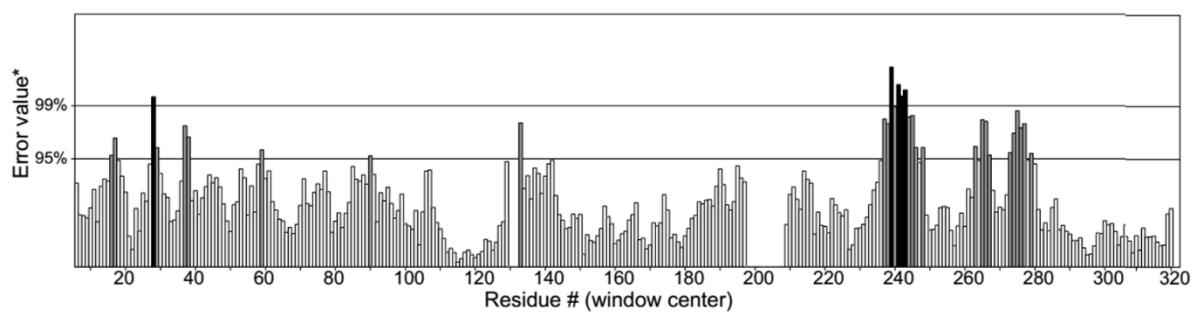


**Fig. S8:** The Verify 3D analysis of *LgBSH*



**Fig. S9:** The overall Z-score values of *LgBSH* analyzed by using ProSA

Program: ERRAT2  
 File: /var/www/SAVES/Jobs/55618009//errat.pdb  
 Chain#:1  
 Overall quality factor\*\*: 90.033



\*On the error axis, two lines are drawn to indicate the confidence with which it is possible to reject regions that exceed that error value.

\*\*Expressed as the percentage of the protein for which the calculated error value falls below the 95% rejection limit. Good high resolution structures generally produce values around 95% or higher. For lower resolutions (2.5 to 3Å) the average overall quality factor is around 91%.

**Fig. S10:** Overall structural quality factor for *LgBSH* evaluated by ERRAT web server.



**Fig. S11:** Substrate binding pocket analysis of *LgBSH* using CASTp online software. Residues involved in the substrate binding includes Cys2, Arg17, Leu19, Asp20, Leu21, Phe25, Lys59, Leu64, Tyr65, Phe66, Gly78, Asn80, Pro99, Phe100, Ile103, Pro104, Gly135, Ser137 and Leu139.